Math 164: Problem Set 3

Owen Jones

1/26/2024

5.3 (a) $f(\mathbf{x}) = (\mathbf{a}^\top \mathbf{x})(\mathbf{b}^\top \mathbf{x})$

$\frac{\partial}{\partial \mathbf{x}_i}(\mathbf{a}^\top \mathbf{x}) = \frac{\partial}{\partial \mathbf{x}_i}\langle \mathbf{a}, \mathbf{x}\rangle = \mathbf{a}_i$

and similarly $\frac{\partial}{\partial \mathbf{x}_i}(\mathbf{b}^\top \mathbf{x}) = \frac{\partial}{\partial \mathbf{x}_i}\langle \mathbf{b}, \mathbf{x}\rangle = \mathbf{b}_i$

so by the product rule $\frac{\partial}{\partial \mathbf{x}_i}(\mathbf{a}^\top \mathbf{x})(\mathbf{b}^\top \mathbf{x}) = (\mathbf{a}^\top \mathbf{x})\mathbf{b}_i + (\mathbf{b}^\top \mathbf{x})\mathbf{a}_i$

$$\Rightarrow \nabla f(\mathbf{x}) = \begin{bmatrix} (\mathbf{a}^\top \mathbf{x})\mathbf{b}_1 + (\mathbf{b}^\top \mathbf{x})\mathbf{a}_1 \\ (\mathbf{a}^\top \mathbf{x})\mathbf{b}_2 + (\mathbf{b}^\top \mathbf{x})\mathbf{a}_2 \\ \vdots \\ (\mathbf{a}^\top \mathbf{x})\mathbf{b}_n + (\mathbf{b}^\top \mathbf{x})\mathbf{a}_n \end{bmatrix}$$

$\Rightarrow \nabla f(\mathbf{x}) = (\mathbf{b}\mathbf{a}^\top + \mathbf{a}\mathbf{b}^\top)\mathbf{x}$

(b) $$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} \frac{\partial^2 f}{\partial \mathbf{x}_1^2} & \frac{\partial^2 f}{\partial \mathbf{x}_2 \partial \mathbf{x}_1} & \cdots & \frac{\partial^2 f}{\partial \mathbf{x}_n \partial \mathbf{x}_1} \\ \frac{\partial^2 f}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} & \frac{\partial^2 f}{\partial \mathbf{x}_2^2} & \cdots & \frac{\partial^2 f}{\partial \mathbf{x}_n \partial \mathbf{x}_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial \mathbf{x}_1 \partial \mathbf{x}_n} & \frac{\partial^2 f}{\partial \mathbf{x}_2 \partial \mathbf{x}_n} & \cdots & \frac{\partial^2 f}{\partial \mathbf{x}_n^2} \end{bmatrix}$$

$\frac{\partial}{\partial \mathbf{x}_i}(\mathbf{a}^\top \mathbf{x})(\mathbf{b}^\top \mathbf{x}) = (\mathbf{a}^\top \mathbf{x})\mathbf{b}_i + (\mathbf{b}^\top \mathbf{x})\mathbf{a}_i \Rightarrow \frac{\partial^2}{\partial \mathbf{x}_i \partial \mathbf{x}_j}(\mathbf{a}^\top \mathbf{x})(\mathbf{b}^\top \mathbf{x})$

$= \mathbf{a}_j \mathbf{b}_i + \mathbf{b}_j \mathbf{a}_i$

$$\Rightarrow \mathbf{F}(\mathbf{x}) = \begin{bmatrix} 2\mathbf{a}_1\mathbf{b}_1 & \mathbf{a}_1\mathbf{b}_2 + \mathbf{a}_2\mathbf{b}_1 & \cdots & \mathbf{a}_1\mathbf{b}_n + \mathbf{a}_n\mathbf{b}_1 \\ \mathbf{a}_2\mathbf{b}_1 + \mathbf{a}_1\mathbf{b}_2 & 2\mathbf{a}_2\mathbf{b}_2 & \cdots & \mathbf{a}_2\mathbf{b}_n + \mathbf{a}_n\mathbf{b}_2 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{a}_n\mathbf{b}_1 + \mathbf{a}_1\mathbf{b}_n & \mathbf{a}_n\mathbf{b}_2 + \mathbf{a}_2\mathbf{b}_n & \cdots & 2\mathbf{a}_n\mathbf{b}_n \end{bmatrix}$$

$= \mathbf{b}\mathbf{a}^\top + \mathbf{a}\mathbf{b}^\top$

5.5 $\frac{\partial}{\partial s}\mathbf{f}(\mathbf{g}(\mathbf{s}, \mathbf{t})) = \frac{\partial \mathbf{f}}{\partial \mathbf{g}_1}\frac{\partial \mathbf{g}_1}{\partial s} + \frac{\partial \mathbf{f}}{\partial \mathbf{g}_2}\frac{\partial \mathbf{g}_2}{\partial s}$

$= \frac{2s+t}{2} \cdot 4 + \frac{4s+3t}{2} \cdot 2 = 8s + 5t$

$\frac{\partial}{\partial t}\mathbf{f}(\mathbf{g}(\mathbf{s}, \mathbf{t})) = \frac{\partial \mathbf{f}}{\partial \mathbf{g}_1}\frac{\partial \mathbf{g}_1}{\partial t} + \frac{\partial \mathbf{f}}{\partial \mathbf{g}_2}\frac{\partial \mathbf{g}_2}{\partial t}$

$= \frac{2s+t}{2} \cdot 3 + \frac{4s+3t}{2} \cdot 1 = 5s + 3t$

5.10 (a) $f(\mathbf{x}) = f(\mathbf{x_0}) + Df(\mathbf{x_0})(\mathbf{x} - \mathbf{x_0}) + \frac{1}{2}(\mathbf{x} - \mathbf{x_0})^\top D^2 f(\mathbf{x_0})(\mathbf{x} - \mathbf{x_0}) + R_3$

$= 1 \cdot e^{-0} + 0 + 1 + [e^{-0}, -1 \cdot e^{-0} + 1] \begin{bmatrix} x_1 - 1 \\ x_2 - 0 \end{bmatrix}$

$+ \frac{1}{2}[x_1 - 1, x_2 - 0] \begin{bmatrix} 0 & -e^{-0} \\ -e^{-0} & 1 \cdot e^{-0} \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 0 \end{bmatrix} + R_3$

$= 2 + (x_1 - 1) + \frac{1}{2}[-x_2, 1 - x_1 + x_2] \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} + R_3$
$= 1 + x_1 + (1 - x_1)(1 - x_2) + \frac{x_2^2}{2} + R_3$

(b) $f(\mathbf{x}) = f(\mathbf{x_0}) + Df(\mathbf{x_0})(\mathbf{x} - \mathbf{x_0}) + \frac{1}{2}(\mathbf{x} - \mathbf{x_0})^\top D^2 f(\mathbf{x_0})(\mathbf{x} - \mathbf{x_0}) + R_3$
$f(\mathbf{x_0}) = 1^4 + 2 \cdot 1^2 \cdot 1^2 + 1^4 = 4$
$f_{x_1}(\mathbf{x_0}) = 4 \cdot 1^3 + 4 \cdot 1 \cdot 1^2 = 8, f_{x_2}(\mathbf{x_0}) = 4 \cdot 1^2 \cdot 1 + 4 \cdot 1^3 = 8$
$f_{x_1x_1}(\mathbf{x_0}) = 12 \cdot 1^2 + 4 \cdot 1^2 = 16, f_{x_1x_2}(\mathbf{x_0}) = 8 \cdot 1 \cdot 1 = 8, f_{x_2x_2}(\mathbf{x_0}) = 4 \cdot 1^2 + 12 \cdot 1^2 = 16$
$f(\mathbf{x}) = 4 + [8, 8] \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix} + [x_1 - 1, x_2 - 1] \begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix} + R_3$
$= 4 + 8(x_1 - 1) + 8(x_2 - 1) + [8(x_1 - 1) + 4(x_2 - 1), 4(x_1 - 1) + 8(x_2 - 1)] \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \end{bmatrix} + R_3$
$= 4 + 8(x_1 - 1) + 8(x_2 - 1) + 8(x_1 - 1)(x_2 - 1) + 8(x_1 - 1)^2 + 8(x_2 - 1)^2 + R_3$
$= 12 + 8x_1^2 + 8x_2^2 - 16x_1 - 16x_2 + 8x_1x_2 + R_3$

(c) $f(\mathbf{x_0}) = e^{1-0} + e^{1+0} + 1 + 0 + 1 = 2e + 2$
$f_{x_1}(\mathbf{x_0}) = e^{1-0} + e^{1+0} + 1 = 2e + 1, f_{x_2}(\mathbf{x_0}) = -e^{1-0} + e^{1+0} + 1 = 1$
$f_{x_1x_1}(\mathbf{x_0}) = e^{1-0} + e^{1+0} = 2e, f_{x_1x_2}(\mathbf{x_0}) = -e^{1-0} + e^{1+0} = 0, f_{x_2x_2}(\mathbf{x_0}) = e^{1-0} + e^{1+0} = 2e$
$\Rightarrow f(\mathbf{x}) = 2e + 2 + [2e + 1, 1] \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} + \frac{1}{2}[x_1 - 1, x_2] \begin{bmatrix} 2e & 0 \\ 0 & 2e \end{bmatrix} \begin{bmatrix} x_1 - 1 \\ x_2 \end{bmatrix} + R_3$
$= 2e + 2 + (2e + 1)(x_1 - 1) + x_2 + e(x_1 - 1)^2 + ex_2^2 + R_3$
$= ex_1^2 + ex_2^2 + x_1 + x_2 + e + 1 + R_3$

6.1 (a) $\mathbf{d}^\top \nabla f(\mathbf{x}^*) = [0, -1] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -1 < 0$ for $\mathbf{d} = [0, -1]^\top$ where $\exists \alpha_0$ s.t $\mathbf{x}^* + \alpha \mathbf{d} \in \Omega \; \forall \alpha \in [0, \alpha_0]$, so by FONC, $\mathbf{x}^*$ is not a local minimum.

(b) $\mathbf{d} = \{[d_1, d_2]^\top : x_1, x_2 \geq 0\}$, so $\mathbf{d}^\top \nabla f(\mathbf{x}^*) = d_1 \geq 0$, so by FONC, $\mathbf{x}^*$ is possibly a local minimum.

(c) $\mathbf{x}^*$ is an interior point of Ω where $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and $\mathbf{F}(\mathbf{x}^*) > 0$, so by SOSC, $\mathbf{x}^*$ is definitely a local minimum.

(d) $\mathbf{d} = \{[d_1, d_2]^\top : x_1, x_2 \geq 0\}$, so $\mathbf{d}^\top \nabla f(\mathbf{x}^*) = [0, 1] \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathbf{0}$. However $\mathbf{d}^\top \mathbf{F}(\mathbf{x}^*) \mathbf{d} = [0, 1] \begin{bmatrix} 0 \\ -1 \end{bmatrix} = -1 < 0$, so by SONC, $\mathbf{x}^*$ is not a local minimum.

6.4 If $\mathbf{x}^*$ is an interior point of Ω, then there exists an open ball $B(\mathbf{x}^*, \delta_1)$ centered at $\mathbf{x}^*$ s.t $B(\mathbf{x}^*, \delta) \subset \Omega$. If $\mathbf{x}^*$ is a local minimizer, then there exists an open ball $B(\mathbf{x}^*, \delta_2)$ centered at $\mathbf{x}^*$ s.t $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in B(\mathbf{x}^*, \delta_2) \cap \Omega$. Pick $\delta = \min\{\delta_1, \delta_2\}$. Because $B(\mathbf{x}^*, \delta) \subset \Omega$ and $\Omega \subset \Omega'$, it follows that $B(\mathbf{x}^*, \delta) \subset \Omega'$. Thus, there exists a neighborhood of values in Ω' s.t $f(\mathbf{x}^*) \leq f(\mathbf{x})$. Hence, $\mathbf{x}^*$ is also a local minimizer over Ω'.

For a counterexample when $\mathbf{x}^*$ is a boundary point:
Consider the function $f(\mathbf{x}) = \mathbf{x}^\top \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \mathbf{x}$ over the region $\Omega = \{\mathbf{x} = [x_1, x_1]^\top : x_1 \geq x_2 \geq 0\}$. $\mathbf{x}^* = [0,0]^\top$ is a boundary point and a local minimizer over Ω, but $\mathbf{x}^*$ is not a local minimizer over $\mathbb{R}^2$.

6.7 Let $\mathbf{y}^* := \arg\min_{\mathbf{y}\in\Omega'} f(\mathbf{y} - \mathbf{x_0})$. It follows $\forall \mathbf{y} \in \Omega'$ $f(\mathbf{y}^* - \mathbf{x_0}) \leq f(\mathbf{y} - \mathbf{x_0})$. If $\mathbf{y}^* \in \Omega' \Rightarrow \mathbf{y}^* - \mathbf{x_0} \in \Omega$. Let $\mathbf{x} = \mathbf{y} - \mathbf{x_0}$. It follows $f(\mathbf{y}^* - \mathbf{x_0}) \leq f(\mathbf{x})$ $\forall \mathbf{x} \in \Omega$ Thus, $\mathbf{y}^* - \mathbf{x_0} = \arg\min_{\mathbf{x}\in\Omega} f(\mathbf{x}) \Rightarrow \mathbf{y}^* = \arg\min_{\mathbf{x}\in\Omega} f(\mathbf{x}) + \mathbf{x_0}$.

6.10 (a) Because $Q = \begin{bmatrix} 2 & 5 \\ -1 & 1 \end{bmatrix}$ is not symmetric, we can replace the matrix with $Q_0 = \frac{1}{2}(Q+Q^\top) = \frac{1}{2}\begin{bmatrix} 4 & 4 \\ 4 & 2 \end{bmatrix}$. It follows $\nabla f(x) = \begin{bmatrix} 4 & 4 \\ 4 & 2 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ and $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} 4 & 4 \\ 4 & 2 \end{bmatrix}$.

Plugging in $\mathbf{d} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ at $\mathbf{x_0} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ we obtain

$$\mathbf{d}^\top \nabla f(\mathbf{x_0}) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^\top \begin{bmatrix} 4 & 4 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}^\top \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 7$$

(b) Want to find all points where $\nabla f(x) = \mathbf{0}$

$\mathbf{x} = -\begin{bmatrix} 4 & 4 \\ 4 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = -\begin{bmatrix} -\frac{1}{4} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -\frac{5}{4} \\ \frac{1}{2} \end{bmatrix}$ is the only point that satisfies the FONC.

The Hessian is not positive semidefinite because the determinant is negative. Thus, $\begin{bmatrix} -\frac{5}{4} \\ \frac{1}{2} \end{bmatrix}$ does not satisfy the SONC. Hence, $\begin{bmatrix} -\frac{5}{4} \\ \frac{1}{2} \end{bmatrix}$ is not a local minimum, so f does not have a minimizer.

6.11 (a) $\nabla f(0) = \begin{bmatrix} 0 \\ -2 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Any vector dotted with $\mathbf{0}$ gives $\mathbf{0}$, so $\mathbf{d}^\top \nabla f(\mathbf{0}) \geq \mathbf{0}$ for all feasible directions.

(b) $[x_1, x_2]^\top = \mathbf{0}$ is a local maximizer because $\mathbf{F}(\mathbf{x}) = \mathbf{x}^\top \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x} \leq 0$ is negative semidefinite. $[x_1, x_2]^\top = \mathbf{0}$ is not strict because if we take $x_1 \neq 0$ and $x_2 = 0$ we obtain $f([x_1, x_2]^\top) = \mathbf{0}$.

6.14 (a) $\nabla f(\mathbf{x}) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ which is nonzero $\forall \mathbf{x} \in \Omega$. Thus, for $\mathbf{x}$ to satisfy the FONC, $\mathbf{x} \in \partial\Omega = \{\mathbf{x} \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$ and $\mathbf{d}^\top \nabla f(\mathbf{x}) = d_2 \geq 0$ for all feasible directions. It follows $[0,1]^\top$ is the only point that satisfies this condition. This can be clearly seen by graphing $x^2 + y^2 = 1$.

(b) $\mathbf{F}(\mathbf{x}) = \mathbf{0}^{2\times 2}$, so $\mathbf{F}(\mathbf{x}) \geq 0$. Thus, any point satisfies the SONC. Hence, $[0,1]^\top$ satisfies the SONC.

(c) $[0,1]^\top$ is not a local minimizer. Consider the set of points $\Omega' = \{\mathbf{x} \in \mathbb{R}^2 : x_1 = \sqrt{1-x_2^2}, x_2 \in [0,1)\}$. Clearly $(\sqrt{1-x_2^2})^2 + x_2^2 = 1 \geq 1 \Rightarrow \Omega' \subset \Omega$. We can make points in Ω as close to $[0,1]^\top$ as we want, but $\forall \mathbf{x} \in \Omega'$ $f(\mathbf{x}) < f([0,1]^\top)$. Hence, f has no minimizer.

6.17 (a) $\nabla f(\mathbf{x}) = [\frac{1}{x_1}, \frac{1}{x_2}]^\top$. Because $\frac{1}{x} \neq 0$ for $x \in \mathbb{R}$, there are no points where $\nabla f(\mathbf{x}) = \mathbf{0}$. Hence, $\mathbf{x}^*$ can't be an interior point.

(b) $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} -\frac{1}{x_1^2} & 0 \\ 0 & -\frac{1}{x_2^2} \end{bmatrix}$ which is negative definite for all x because $-\mathbf{F}(\mathbf{x})$ is positive definite. Thus, every $\mathbf{x}$ satisfies the SONC.